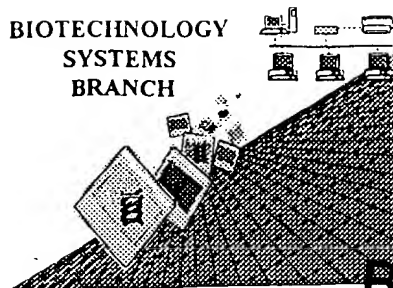


0

RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY
SYSTEMS
BRANCH



RECEIVED

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

JUL 06 2001
OFFICE OF PETITIONS
DEPUTY A/C PATENTS

Application Serial Number: 08/786,533

Source: 1642

Date Processed by STIC: 6-12-01

RECEIVED

JUL 05 2001

TECH CENTER 1600/2900

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST 25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

RECEIVED

JUL 05 2001

Raw Sequence Listing Error Summary

TECH CENTER 1600/2900

ERROR DETECTED SUGGESTED CORRECTIONSERIAL NUMBER: 08/786,533

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☐ **Wrapped Nucleics
Wrapped Aminos** The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 ☐ **Invalid Line Length** The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ☐ **Misaligned Amino
Numbering** The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 ☐ **Non-ASCII** The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ☐ **Variable Length** Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 ☐ **PatentIn 2.0
"bug"** A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 ☐ **Skipped Sequences
(OLD RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 ☐ **Skipped Sequences
(NEW RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 9 ☒ **Use of n's or Xaa's
(NEW RULES)** Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ☐ **Invalid <213>
Response** Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ☐ **Use of <220>** Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 ☐ **PatentIn 2.0
"bug"** Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

1642

RAW SEQUENCE LISTING

DATE: 06/12/2001

PATENT APPLICATION: US/08/786,533

TIME: 13:24:37

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06122001\H786533.raw

Does Not Comply
Corrected Diskette Needed

See pp. 1-4

4 <110> APPLICANT: Horwitz, Marcus A.
 5 Harth, Gunter
 6 Lee, Bai-Yu
 8 <120> TITLE OF INVENTION: Abundant Extracellular Products and
 9 Methods for Their Production and Use
 12 <130> FILE REFERENCE: 510030-143
 14 <140> CURRENT APPLICATION NUMBER: US 08/786,533
 15 <141> CURRENT FILING DATE: 1997-01-21
 17 <150> PRIOR APPLICATION NUMBER: US 08/568,357
 18 <151> PRIOR FILING DATE: 1995-12-06
 20 <150> PRIOR APPLICATION NUMBER: US 08/551,149
 21 <151> PRIOR FILING DATE: 1995-10-31
 23 <150> PRIOR APPLICATION NUMBER: US 08/447,398
 24 <151> PRIOR FILING DATE: 1995-05-23
 26 <150> PRIOR APPLICATION NUMBER: US 08/289,667
 27 <151> PRIOR FILING DATE: 1994-08-12
 29 <150> PRIOR APPLICATION NUMBER: US 08/156,358
 30 <151> PRIOR FILING DATE: 1993-11-23
 32 <160> NUMBER OF SEQ ID NOS: 160
 34 <170> SOFTWARE: FastSEQ for Windows Version 3.0

ERRORED SEQUENCES

234 <210> SEQ ID NO: 21
 235 <211> LENGTH: 47
 236 <212> TYPE: PRT
 237 <213> ORGANISM: Mycobacterium tuberculosis
 239 <400> SEQUENCE: 21
 240 Ala Tyr Pro Ile Thr Gly Cys Leu Gly Ser Glu Leu Thr Met Thr Asp
 241 1 5 10 15
 242 Thr Val Gly Gln Val Val Leu Gly Trp Lys Val Ser Asp Leu Phe Lys
 243 20 25 30
 E--> 244 Ser Thr Ala Val Ile Pro Gly Tyr Thr Val Xaa Glu Gln Gln Ile
 245 35 40 45
 247 <210> SEQ ID NO: 22
 248 <211> LENGTH: 47
 249 <212> TYPE: PRT
 250 <213> ORGANISM: Mycobacterium tuberculosis
 252 <400> SEQUENCE: 22
 253 Ala Tyr Pro Ile Thr Asx Lys Leu Gly Ser Glu Leu Thr Met Thr Asp
 254 1 5 10 15
 255 Thr Val Gly Gln Val Val Leu Gly Trp Lys Val Ser Asp Leu Tyr Lys
 256 20 25 30
 E--> 257 Ser Thr Ala Val Ile Pro Gly Tyr Thr Val Xaa Glu Gln Gln Ile
 258 35 40 45
 271 <210> SEQ ID NO: 24

Missing mandatory <220> to <223>
 features to explain the Xaa in
 the sequence. See # 9 on the

Error Summary
 Sheet.

see above

RAW SEQUENCE LISTING

DATE: 06/12/2001

PATENT APPLICATION: US/08/786,533

TIME: 13:24:37

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06122001\H786533.raw

```

272 <211> LENGTH: 13
273 <212> TYPE: PRT
274 <213> ORGANISM: Mycobacterium tuberculosis
276 <400> SEQUENCE: 24
E--> 277  Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp
      278    1          5          10
280 <210> SEQ ID NO: 25
281 <211> LENGTH: 60
282 <212> TYPE: PRT
283 <213> ORGANISM: Mycobacterium tuberculosis
285 <400> SEQUENCE: 25
286  Ala Pro Tyr Glu Asn Leu Met Asx Pro Ser Pro Ser Met Gly Arg Asp
287    1          5          10          15
288  Ile Pro Val Ala Phe Leu Ala Gly Gly Pro His Ala Val Tyr Leu Leu
289        20          25          30
290  Asp Ala Phe Asn Ala Gly Pro Asp Val Ser Asn Trp Val Thr Ala Gly
291      35          40          45
E--> 292  Asn Ala Met Met Thr Leu Ala Xaa Lys Gly Ile Cys
      293    50          55          60
295 <210> SEQ ID NO: 26
296 <211> LENGTH: 60
297 <212> TYPE: PRT
298 <213> ORGANISM: Mycobacterium tuberculosis
300 <400> SEQUENCE: 26
301  Ala Pro Tyr Glu Asn Leu Met Val Pro Ser Pro Ser Met Gly Arg Asp
302    1          5          10          15
303  Ile Pro Val Ala Phe Leu Ala Gly Gly Pro His Ala Val Tyr Leu Leu
304        20          25          30
305  Asp Ala Phe Asn Ala Gly Pro Asp Val Ser Asn Trp Val Thr Ala Gly
306      35          40          45
E--> 307  Asn Ala Met Met Thr Leu Ala Xaa Lys Gly Ile Ser
      308    50          55          60
323 <210> SEQ ID NO: 28
324 <211> LENGTH: 40
325 <212> TYPE: PRT
326 <213> ORGANISM: Mycobacterium tuberculosis
328 <400> SEQUENCE: 28
329  Phe Ser Arg Pro Gly Leu Pro Val Glu Tyr Leu Gln Val Pro Ser Pro
330    1          5          10          15
331  Ser Met Gly Arg Asp Ile Lys Val Gln Phe Gln Ser Gly Gly Asn Asn
332      20          25          30
E--> 333  Ser Pro Xaa Leu Tyr Leu Leu Asp
      334      35          40
336 <210> SEQ ID NO: 29
337 <211> LENGTH: 22
338 <212> TYPE: PRT
339 <213> ORGANISM: Mycobacterium tuberculosis
341 <400> SEQUENCE: 29
342  Phe Ser Arg Pro Gly Leu Pro Val Glu Tyr Leu Gln Val Pro Ser Ala

```

Same

Same

Same

Same

RAW SEQUENCE LISTING

DATE: 06/12/2001

PATENT APPLICATION: US/08/786,533

TIME: 13:24:37

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06122001\H786533.raw

```
343      1      5      10      15
E--> 344  Xaa Met Gly Arg Asp Ile
345      20
347 <210> SEQ ID NO: 30
348 <211> LENGTH: 30
349 <212> TYPE: PRT
350 <213> ORGANISM: Mycobacterium tuberculosis
352 <400> SEQUENCE: 30
353  Asp Pro Glu Pro Ala Pro Pro Val Pro Asp Asp Ala Ala Ser Pro Pro
354      1      5      10      15
E--> 355  Asp Asp Ala Ala Ala Pro Pro Ala Pro Ala Asp Pro Pro Xaa
356      20      25      30
387 <210> SEQ ID NO: 34
388 <211> LENGTH: 22
389 <212> TYPE: PRT
390 <213> ORGANISM: Mycobacterium tuberculosis
392 <400> SEQUENCE: 34
393  Asn Ser Lys Ser Val Asn Ser Phe Gly Ala His Asp Thr Leu Lys Val
394      1      5      10      15
E--> 395  Xaa Glu Arg Lys Arg Gln
396      20
```

Same

Same

Same

VERIFICATION SUMMARY

DATE: 06/12/2001

PATENT APPLICATION: US/08/786,533

TIME: 13:24:39

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06122001\H786533.raw

L:244 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:21
L:257 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:22
L:277 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:24
L:292 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:25
L:307 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:26
L:333 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:28
L:344 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:29
L:355 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:30
L:395 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:34